

# Processamento de Fala 2007/08

## 3º Mini-teste

14 de Dezembro de 2007

Identifique o seu enunciado colocando o seu nome e número de aluno no espaço reservado no final. Só são aceites respostas às questões de escolha múltipla assinaladas no local apropriado no final do enunciado quando este estiver identificado.

1. Diga se as seguintes afirmações são verdadeiras ou falsas:

- (a) Num sistema de reconhecimento baseado em modelos de palavras, é fácil de modelar a coarticulação entre elas.
- (b) Os modelos sub-palavra têm frequentemente 3 estados sem salto sentre estados de modo a introduzir restrições de duração mínima.
- (c) Adoptam-se frequentemente matrizes diagonais de forma a restringir o número de parâmetros a serem treinados.
- (d) A taxa de erro de reconhecimento aumenta proporcionalmente com o quadrado da perplexidade.
- (e) Em descodificação multi-passo, o primeiro passo usa tipicamente modelos de trifones.
- (f) As medidas de confiança podem ser usadas para adaptação ao orador não supervisionada.
- (g) Na razão de máxima verosimilhança o denominador é a energia do resíduo resultante de filtrar a locução de teste por cada um dos filtros inversos de referência.
- (h) Os modelos HMM ergódicos exibem estrutura temporal.

2. Como se designam os modelos HMM que partilham o mesmo conjunto de distribuições Gaussianas com diferentes pesos?

3. Como se designam os modelos de língua estocásticos em que, para cada palavra no vocabulário, a gramática fornece uma lista de palavras que lhe pode suceder, mas todas as palavras nessa lista são igualmente prováveis?

4. Qual a taxa de erro de palavras de um humano para um vocabulário entre 20k-30k palavras isoladas , sem informação sintáctica, nem semântica?

5. O algoritmo de Viterbi é frequentemente usado para resolver dois dos problemas básicos em HMMs. Indique quais.

6. Indique a expressão matemática para o cálculo dos parâmetros delta-delta cepstrais, em função dos delta-cepstrais.

7. Dê exemplos de duas palavras, indicando a sua transição fonética em situações de pronúncia canónica ou pouco cuidada.

- (a) A primeira palavra deve ter uma sílaba sem coda, em que o núcleo muda em pronúncia pouco cuidada;
- (b) A segunda palavra deve ter uma sílaba com ataque núcleo e coda, em que esta última cai em pronúncia pouco cuidada.

Nota: sublinha a sílaba pedida em cada uma das palavras.

8. Nas aulas de laboratório foi feita uma demo de um sistema de reconhecimento de palavras isoladas (meses) dependente do orador. Nessa demo, os parâmetros extraídos pela ferramenta HCOPY (do pacote HTK) tinham o formato “MFCC\_E\_D\_Z”. Diga por extenso qual o significado de cada uma das 4 partes.

9. Considere um sistema de reconhecimento de fala contínua e vocabulário extenso, independente do orador. Para um dado segmento de fala, com algum ruído de fundo, a transcrição manual feita foi a seguinte:

*O Futebol Clube do Porto voltou às vitórias.*

*O líder do campeonato bateu o Vitória de Setúbal por dois zero.*

*Na noite fria do Dragão não foi preciso esperar mais que cinco minutos para ver Lisandro Lopez fazer o golo dez da conta pessoal e colocar o Porto a ganhar por um zero.*

*A primeira parte teve tons azuis e ao quarto de hora o avançado argentino mostrou que o bom momento de forma é mais do que apenas marcar.*

A transcrição automática produzida pelo reconhecedor foi:

*O futebol do Porto voltou às vitórias.*

*O líder do campeonato bateu o Vitória de Setúbal por dois zero.*

*Na noite fria do Dragão não foi preciso esperar mais cinco anos para ver lisandro Lopez fazer golo desta conta pessoal colocar o Porto ganhar por um a zero.*

*Na primeira parte esteve doze anos e um quarto de hora o avançado argentino mostrou que de momento de forma é mais do que apenas marcar.*

Ignorando a pontuação e capitalização, complete os valores de H ("correct"), D ("deletions"), S ("substitutions"), I ("insertions"), N ("total"), %Corr, %Acc e %WER correspondentes.

Admita que cada palavra tem associada uma característica de capitalização *sem\_maiúscula / com\_maiúscula*. Calcule a exactidão (%ACC) do capitalizador automático. Não contabilize como erros de capitalização as palavras cuja maiúscula inicial se deverem à pontuação que as antecede. *Transporte* para a transcrição automática a capitalização correcta e faça os cálculos sobre a transcrição automática.

10. Considere o corpus de treino composto pelas seguintes frases:

*A Ana gosta de bolo de chocolate.*

*A Teresa prefere mousse de chocolate.*

*A Joana gosta mais de bolo de manga.*

*A Ana não gosta de mousse de manga.*

*A Teresa não gosta de gelado.*

Considere a frase de teste:

*A Joana gosta de gelado de manga.*

- Calcule o número de unigramas, bigramas e trigramas do corpus de treino, e a dimensão do vocabulário.
- Calcule a probabilidade da frase de teste usando um modelo de bigramas sem alisamento.
- Calcule a probabilidade da frase de teste usando um modelo de bigramas com alisamento do tipo add-one.
- Construa uma frase o mais longa possível (em termos de  $n^\circ$  de palavras) com uma probabilidade de trigramas não nula, mas diferente das existentes.
- Existe algum quadrigrama no corpus de treino com mais que uma ocorrência? Se sim, indique-o.

(Nota: Indique apenas os cálculos a partir dos valores numéricos)

**Respostas**

Nome:	
Número:	

1. (2,4 val.)

a	b	c	d	e	f	g	h

2 a 6 (1 / 1 / 1 / 1,4 / 1,6 valores)

2	
3	
4	
5	
6	

7. (1,8 val.)

palavra	canónica	pouco cuidada

8. (2 val.)

MFCC	
E	
D	
Z	

9. (3,5 val.)

H	D	S	I	N	% Corr	% Acc	% WER	% Acc-Cap

10. (4,3 val.)

a)	
b)	
c)	
d)	
e)	